1.0

1.1

1.25   1.4

4.5
5.0
5.6

2.8   2.5

3.2   2.2

3.6

4.0   2.0

1.8

1.6

MICROCOPY RESOLUTION TEST CHART

LEVEL

QUANTIZATION ERROR DUE TO SIGNATURE TABLE EVALUATIONS

AFIT/GST/EE/80-1

Master's thesis

DAVID S. PRAHLER
USAF

Mar 80

80

AFIT/GST/EE/80-1

QUANTIZATION ERROR DUE TO SIGNATURE

TABLE EVALUATIONS

THESIS

Presented to the Faculty of the School of Engineering

of the Air Force Institute of Technology

Air Training Command

in Partial Fulfillment of the

Requirements for the Degree of

Master of Science

by

DAVID S. PRAHLER

Capt   USAF

Graduate Strategic and Tactical Sciences

March 1980

Approved for public release; distribution unlimited.

Preface

This particular topic resulted from many months of "wandering" in the area of Command, Control, and Communications. My sincere thanks go to my thesis advisor, Major Joe Carl, for his expert guidance in its selection. He possesses that ability to push a student in the right direction without letting it feel like a shove. A special note of appreciation must go to Mr. Royce Reiss at AF ACS/SA at the Pentagon for withstanding many long and sometimes confusing telephone conversations. Lastly, but in my mind, always first, my thanks go to my wife Mary and my children, Adrienne and Christopher, for their unmatched love, support, and patience throughout this ordeal.

DAVID S. PRAHLER

TABLE OF CONTENTS

List of Figures

List of Tables

AFIT/GST/EE/80-1

Abstract

The signature tables used in TAC ASSESSOR were assumed
to be representations of some corresponding value functions
evaluated at certain levels of the input features. The
theoretical mean-squared-error due to this substitution was
calculated for signature table 1 and table 5 in the FAC
Target Importance Tables. The calculations were performed
under two different assumption concerning the assumed value
functions.

Case 1 was that the signature table values resulted from
evaluating the assumed value function at the given input
feature values. Case 2 was that the function was evaluated
at some feature values in the given quantization level
indicated by the input feature. The mean squared error was
then expressed as a percentage of the assumed mean squared
signal. The percent errors derived were 20.48% and 27.81%
for signature table 1 under Case 1 and 2 assumptions, respectively.

Next, an evaluation simulation was performed on three sets
of 30 targets whose features were generated using uniform,
exponential, and normal random variate generators scaled to
the appropriate feature limits. The major result of this
was that a significant number of the targets were involved
in ties when evaluated using the signature table method. The
number of ties ranged from 13 out of 30 up to 28 out of 30
targets involved in evaluation ties.

Finally, the Spearman rank correlation coefficient was used to test the null hypothesis that there was no association between the evaluation methods. The null hypothesis was rejected in every case at the 0.99 confidence level. Based on these results, it is the recommendation of this study that formal application of multi attribute utility assessment techniques be applied to generate a substitute for the signature table evaluation technique applied in this case.

I. Introduction

The question of Command and Control (C^2) System effectiveness has become increasingly important to the military establishment. Some of the reasons for this increased importance include the increasing value of limited resources and the possible consequences of wrong decisions due to inadequate, inaccurrate, and erroneous C^2 information. As resources become more scarce and valuable, it becomes imperative to increase their effective employment. As the magnitude of the possible consequences of s ginale decision increase, so must the amount of invalid C^2 information decrease.

As the costs of research, development, procurement, and maintenance of elements of the C^2 system increase, experimentation with modifications and new equipment has a tendency to decrease in effectiveness. The military establishment and, in particular, the Air Force, has moved towards system simulation to evaluation proposed changes to C^2 systems. The common goal of this movement is to gain a better understanding of the relationship between C^2 and force effectiveness.

A major stumbling block in C^2 simulation has been and continues to be the simulation of the decision making process of commanders within the system. One alternate approach to simulating the decision makers is to employ interactive or "man in the loop" schemes. There are, however, some serious drawbacks to this approach. The first is the time and cost involved with gathering together suitable personnel for valid

1

system operation. Second, is the time required to execute completely interactive model. Finally, the lack of consistency in the simulations due to different players may make it difficult to verify or validate the results. These deficiencies have generated a movement towards the field of artificial intelligence in order to simulate the decision processes necessary.

Artificial intelligence techniques attempt to capture the decision making process to a certain degree so that a given input will yield a consistent response from the simulation. Signature table analysis is one of these techniques and has been applied in C^2 system modeling as a means of simulating the evaluation of situations, alternatives, and targets.

Research Question

The application of signature table analysis to the evaluation of some class of objects requires that the class of objects be characterized in terms of features. These features must be quantized for evaluation by the signature table and this quantization process generates error in the evaluation. What error results from the application of signature evaluation techniques to objects which are characterized in terms of continuous features which must be quantized?

Methodology

The first step in this study will be the development of a theoretical measure of error. This error measure will be derived by assuming that a signature table is a piecewise representation of a continuous value function derived from the decision maker's preferences evaluated over the given features.

2

This theoretical error measure will be converted to a percentage error measure for intuitive appeal and will be applied to the FAC Target Importance Tables used in the TAC ASSESSOR C^2 system model under development at Air Force Studies and Analysis. Finally, sample target data sets will be generated and evaluated using both the signature table evaluation scheme and the assumed value function method. The targets will be rank ordered and the Spearman rank correlation coefficient will be used to test the null hypothesis that a significant difference exists between the two evaluation schemes.

Overview

A description of signature table evaluation techniques is presented in Chapter II, and Chapter III contains a discussion of how these techniques are applied in the TAC ASSESSOR model. Chapter IV develops the error measures to be used in Chapter V on the FAC Target Importance Tables used in TAC ASSESSOR. The last chapter contains the conclusions and recommendations of the study.

II. Signature Tables

Signature tables as discussed in this study are a means of evaluating some object or situation. The resultant evaluation is a measure of worth or utility that represents the decision maker's preference for the combination of features that represent the object being evaluated.

The first step in the application of signature tables as an evaluation tool is to determine a set of features that will characterize that class of objects. For example, a class U of ICBM's might be characterized by a set of n features $\{F_1, F_2, F_3, \ldots, F_n\}$ where

$$F_1 = \text{payload capacity}$$

$$F_2 = \text{range}$$

$$F_3 = \text{reliability}$$

$$F_4 =$$

$$\cdot$$

$$\cdot \qquad\qquad \text{other characteristics}$$

$$\cdot$$

$$F_n =$$

Each object to be evaluated is then measured in terms of the features to be used in the evaluation. These measured values are then quantized for use by the signature table evaluation scheme.

Feature Quantization

Since the signature table can accept only a finite number of input points, it must be decided how many different values each feature will be allowed to assume. The number of quantization levels, q_i, determines the range of values that feature F_i can assume. The number of quantization levels will depend on how unique each object is when measured by the ith feature. There must also be determined a set of threshold values which serve to assign the measurements made into the appropriate quantization level. Finally, the actual values which will represent the quantization level of a particular object being measured have to be determined. The most common scheme is to assign the middle quantization level the value 0 and assign +1, +2, . . . to quantization levels above and -1, -2, . . . to quantization levels below.

In the example, suppose the largest ICBM payload capacity in the class U is 10 Megatons (10MT) and the smallest is 50 Kilotons (50KT). Assume for this example that only three quantization levels are needed to preserve the information contained in the payload capacity feature. One explanation for this may be that the payloads need only be characterized in terms of low, medium, or high. Another could be that the actual payloads naturally divide into three groups. The threshold values may be derived from natural divisions in the data if present or assigned on some other basis. For this example let the first threshold value be 500KT and the second 1MT. Any payload less than 500KT will be assigned the first quantization level. A payload between 500KT and

5

1MT will be assigned the second, and the third level will be
assigned to payloads greater than 1MT. Finally, quantized
values need to be assigned to the quantization levels. For
the purposes of this example let -1, 0, +1 be assigned to
the first, second, and third levels, respectively. The
desired result is that each feature has assigned to it the
smallest possible number of quantization levels in order to
keep the size of the signature tables manageable.

Signature Type Formation

The next step in signature table evaluation is to group
features into subsets. These subsets of features are known
as signature types (Ref 5:610). No generally accepted theory
exists on the method of grouping the features into signature
types and the question that presents itself is whether to
group highly correlated features into signature types or
spread these correlated features among the set of signature
types. (Ref 3:77). According to Page, both combinations
can yield "excellent" tables (Ref 3:82).

The last step in the table construction is the determination
of a derived feature for each possible combination of values
in each signature type. For example, there are three features
in a particular signature type $\{F_1, F_3, F_4\}$ and each of these
features has q_i quantization levels, then there are
$(q_1) (q_3) (q_4)$ or $\prod_i (q_i)$, i = 1,3,4 combinations of values
possible in that signature type. Each of these possible
combinations is a signature.

Continuing with the ICBM evaluation example, let features
F_1, F_3, and F_4 be grouped into one such signature type.

6

Suppose the quantization levels have been chosen as follows:

$q_1 = 3 \ (-1, 0, +1)$

$q_3 = 2 \ (0, +1)$

$q_4 = 3 \ (-1, 0, +1)$

The signature table for this signature type would be
constructed with the tabular headings as indicated in Figure 1.
For each of the 18 possible signatures a derived feature value
is determined which represents an evaluation of each signature.

Derived Feature Determination

Determination of the value of the derived feature for
each signature can be accomplished using several approaches.
These approaches, however, may be grouped into two general
strategies. The first is simply a direct assessment technique.
The analyst or data user determines the relative merit of each
signature and assigns some value D to each of the $\prod_i (q_i) = k$
possible signatures for each signature type. The other
general method of derived feature value determination involves
the use of a computer algorithm to make the assignment based on
analysis of sample data. This procedure is often referred to as
training. In order to train a signature table, a set of input data is
generated. Each object in the training sample is measured in
terms of the n features previously selected and also is assigned
a key feature, K, which characterizes the object in general
terms such as good/bad or desirable/undesirable. At this point,
the mth object of the training sample can be described by

SIGNATURE			DERIVED FEATURE
F_1	F_3	F_4	D
-1	0	-1	
-1	0	0	
-1	0	+1	
-1	1	-1	
-1	1	0	
-1	1	1	
0	0	-1	
0	0	0	
0	0	1	
0	1	-1	
0	1	0	
0	1	1	
1	0	-1	
1	0	0	
1	0	1	
1	1	-1	
1	1	0	
1	1	1	

Fig 1. Example Signature Table Organization

a data set, S_m, which consists of values, $X_{i,m}$, assigned to each feature, F_i, and a key feature value, K_m.

$$S_m = \{X_{1m}, X_{2m}, X_{3m}, \ldots, X_{nm}, K_m\}$$

All data sets of the training sample comprise the input to a computer algorithm that determines the value of D for each of the k possible signatures in each signature type. Each object's data set is checked for the values of the features for each signature type. For the ICBM example, let one ICBM's data set be

$$S_m = \{0, 3, 1, -1, \ldots, K = 1\}$$

The representative signature table of Figure 1 uses a signature type composed of $\{F_1, F_3, F_4\}$ and the corresponding values of that signature type for the mth ICBM taken from S_m would be $\{0, +1, -1\}$. The signature table is searched for this particular signature and a count is made of the key feature values of that signature. The derived feature value, D, corresponding to that signature is computed based on the number of times that particular signature came from a training object with a good key feature (K=1) versus the number of times it came with a bad key feature (K=0). Since each signature type consists of only a few of the features used to characterize the objects, the same signature could result from an object in the training sample with either a good or bad key feature.

The resulting derived features from each signature table comprise a new derived data set for each object. These

derived features can be grouped in a similar manner and higher level signature tables are the result. The final output from such a hierarchy of signature tables depends on the original data in a non-linear way (Ref 3:81).

Current Applications

A. L. Samuel has done considerable work in this area by applying signature table methods to a checkers playing computer program. He characterized moves in the game by 24 features which described the resultant board situation. The 24 features were divided into six signature types, each consisting of four features. His training sample consisted of a library of master play which described approximately 250,000 board situations. The master library games were played by the computer during training and each possible move (object) was evaluated in terms of the 24 features. Each move was either a move recommended by the master library (good key feature) or it was not recommended (bad key feature) given the current board situation. The signature tables were entered for each signature from each move together with the key feature value. Two totals, A and B, were tabulated for each of the possible signatures of each signature type. Each move that was not recommended by the master library when entered into a signature table caused an increase, by one, in the B total of that particular signature. A move recommended by the master library when entered into a signature table caused an increase, by n, in the A total of that particular signature, where n was the number of non-recommended moves in the current situation. Samuel's rationale

10

for this particular scheme was to place more importance on the master library recommended moves and preserve the zero-sum nature of the game (Ref 5:613). Once the training was complete, values for the derived feature of each signature were determined by calculation of a number D for each signature, where

$$D = \frac{A-B}{A+B} \tag{1}$$

If the output of this table, the D value, was the final evaluation in a hierarchy, then the scheme was complete and the value of D which resulted was used in the decision making process. If, however, this signature table was an intermediate table in a hierarchy, the D values had to be quantized for input to the next table. The D values were quantized by first rank ordering the values and then dividing the list into the desired number of quantization levels and quantized values assigned to each group (Ref 5:613). If five quantization levels wes desired, the rank ordered D values would be divided into five equal sized groups. The signatures which yielded D values in the highest group would be assigned a +2, those in the lowest group would be assigned a -2. Values of +1, -1, and 0 would then be assigned to those signatures which yielded calculated D values which fell into the second highest, second lowest, and middle groups, respectively.

C. V. Page has also applied the concept of signature table analysis. He used the technique to evaluate housing

quality in Detroit, Michigan, on the basis of 12 features.
Some of the 12 features he used were rubble per vacant
lot, dwelling units per house, and trash per dwelling unit
(Ref 3:78). Organization of his data was similar to Samuel's
application; however, he added a statistical test for signi-
ance to the value assessment of the derived feature in
each table. Page assumed that an estimate of the key
feature of a data set being high was known (p). Next, a
confidence interval was chosen such that $C = 1-\alpha$. Using p
and C, he calculated the critical values N_1 and N_2 from the
binomial distribution such that:

$$P(A \leq N_1-1) < \frac{C}{2} \leq P(A \leq N_1) \quad , \tag{2}$$

and

$$P(A \leq N_2-1) < 1-\frac{C}{2} \leq P(A \leq N_2) \tag{3}$$

where A was a different count mechanism than it was in
Samuel's scheme. Page used A to represent the number of times
a particular signature was the result of a training subject
with a high key feature. The values of N_1 and N_2 were the
calculated threshold values in a three level scheme. The
value of the derived feature for each signature was assigned
on the basis of A and its relationship to the threshold
values. If $A > N_2$, D for that signature was assigned the
high value, +1. Low value, -1, was assigned if $A \leq N_1$ and
0 was assigned if $N_1 < A \leq N_2$. Page called this the interval
method. It resulted in the assignment of +1 and -1 to

"unusual" events and 0 to events which could be explained by chance (Ref 3:80).

Using a single confidence interval, Page's method yielded three values that the derived feature could be assigned. Using two confidence intervals, the number of values available can be increased to five (Ref 3:80). In this case, threshold values N_1, N_2, N_3, and N_4 would be calculated such that the derived feature, D, took the following values for the given conditions:

$$D = -2, \quad N_1 \geq A \tag{4}$$

$$D = -1, \quad N_1 < A \leq N_2 \tag{5}$$

$$D = 0, \quad N_2 < A \leq N_3 \tag{6}$$

$$D = +1, \quad N_3 < A \leq N_4 \tag{7}$$

$$D = +2, \quad N_4 < A \tag{8}$$

Both Page and Samuel used a training sample in order to calculate the derived feature for each signature in each table. From these derived features, second level tables were generated. The derived features were grouped into signature types in the same manner as those generated for the first level. The second level tables yielded second level derived features which could, in turn, be used to generate higher level tables. Samuel used, among others, a three level hierarchy with 24 entry features grouped into six signatures types at the first level. This hierarchy

is shown in Figure 2 (Ref 5:611). The organization used by Page is shown in Figure 3 (Ref 3:84). Twelve entry features were used to construct 66 first level tables which correspond to all possible pairwise combinations of the entry features. As can be seen, each entry level feature is used in more than one signature type and table. These redundancies, with redundancies in the grouping of higher level derived features, yielded a six level hierarchy.

The end result of the direct assessment or training methods is the same: A hierarchical structure of tables, each consisting of a list of possible signatures and a resultant derived feature value for each signature. Once the derived features for all signatures have been determined, evaluation of new or unassessed objects can be made. The tables now act as a predictor: They makes inferences about new objects submitted for analysis. Given a new signature, does that signature come from an object that is good or bad? The final derived feature is the answer to that question.
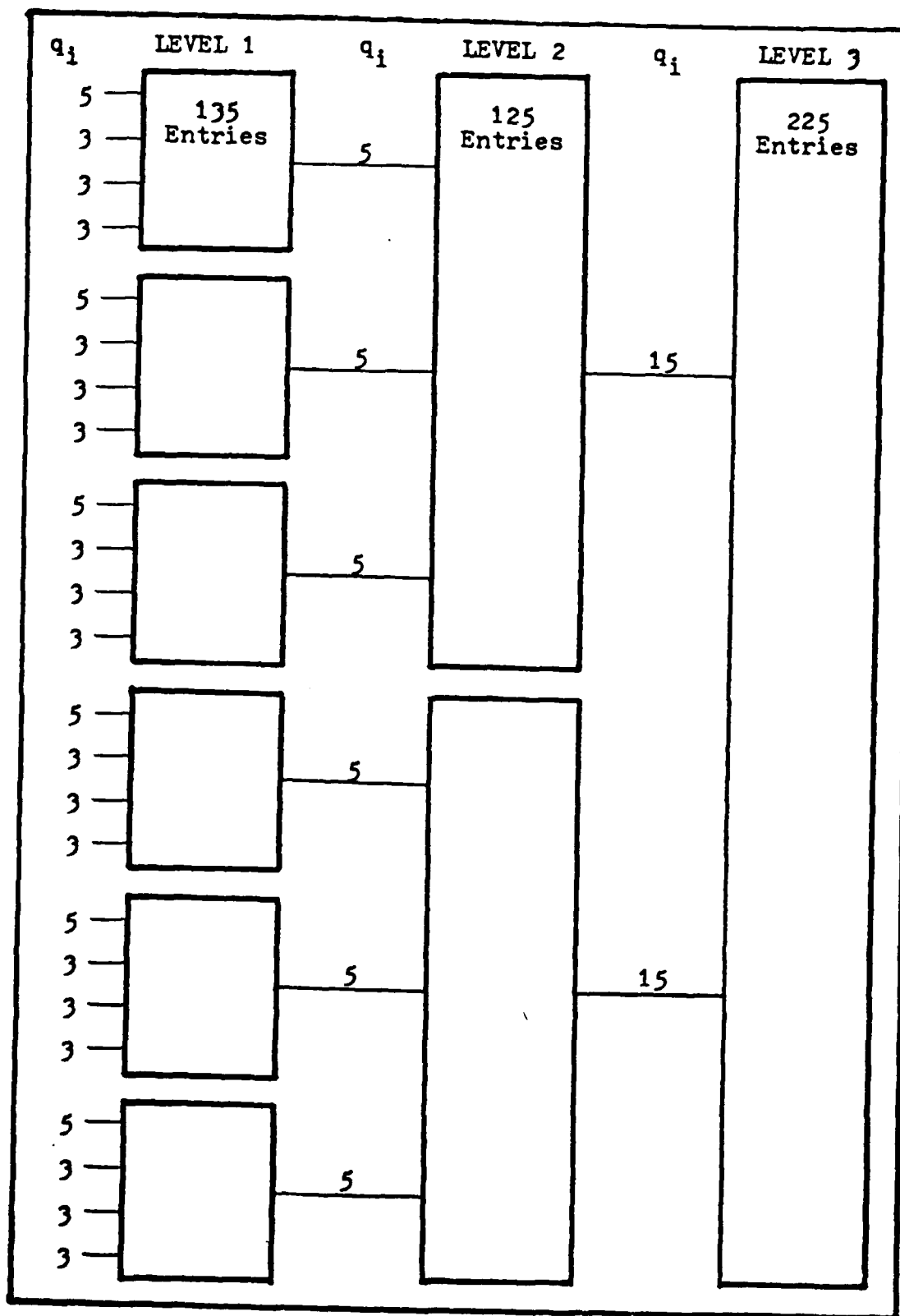
Fig 2. Signature Table Hierarchy Used by Samuel

15
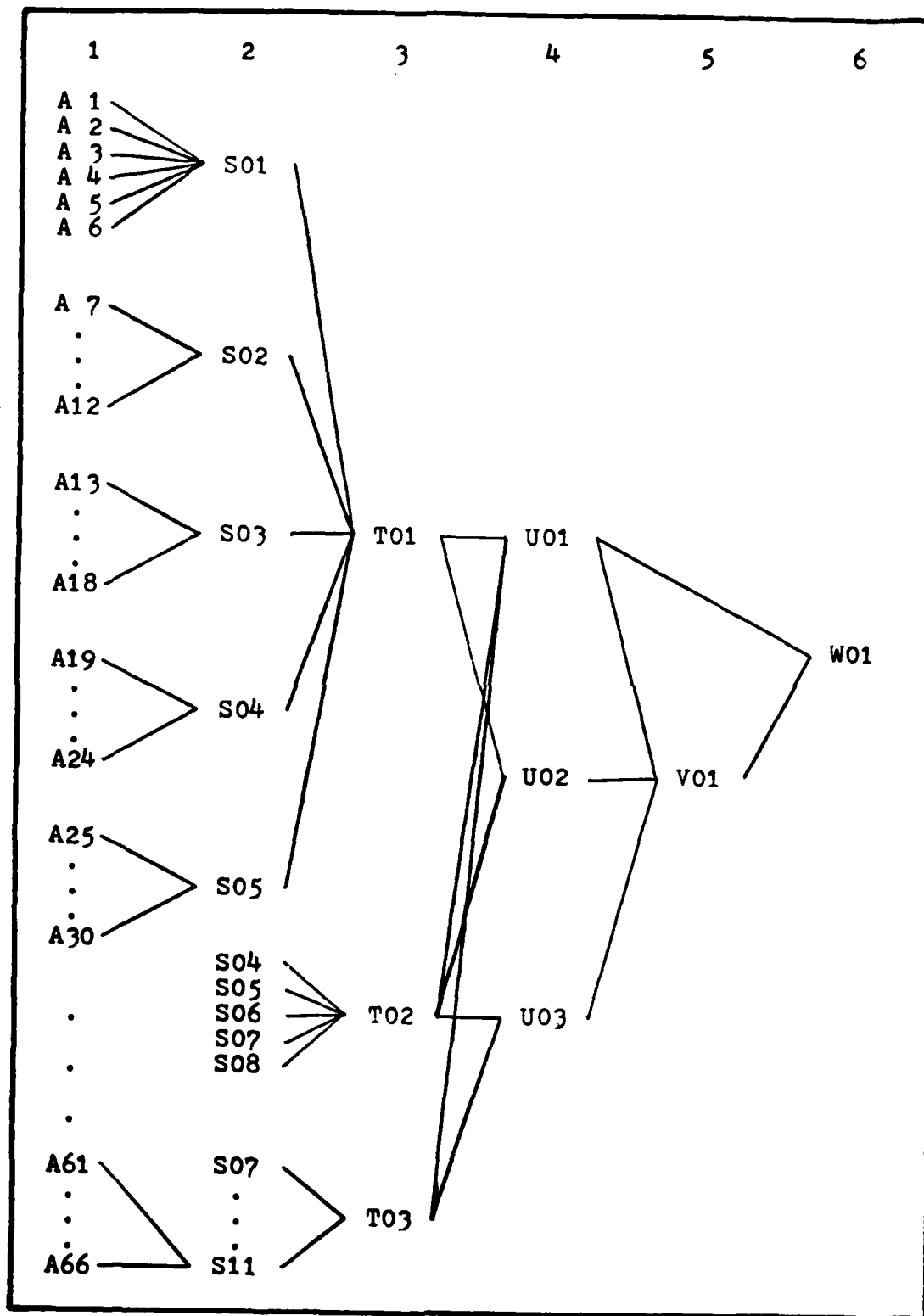
Fig 3. Signature Table Hierarchy Used by Page

16

III. Signature Tables in TAC ASSESSOR

The use of signature tables in the TAC ASSESSOR combined-arms simulation model is but part of the overall effort to simulate the decision process of individual commanders and staffs in a theater level conflict. The emphasis is on the simulation of the decision maker (DM) who distills large sets of data and extracts important information from a situation in terms of combinations of features which characterize the situation. In this respect, a signature table "mimics a skilled analyst who, faced with the task of analyzing an extremely large and complex set of data, is forced to restrict his attention to those data items or combinations of data items that are in some sense significant" (Ref 6:61). This chapter will focus on the uses of signature tables in TAC ASSESSOR and a detailed look at a particular hierarchy of tables used in the ranking of targets against which close air support (CAS) sorties are assigned.

Close air support target ranking is one of five processes supported through the use of signature tables. The signature table technique, along with production systems, heuristics for decision tree analysis, and directed relational graphs comprise the set of artificial intelligence techniques used in the model. Signature table techniques are applied specifically to the processes of target correlation, battle assessment, plan evaluation and sortie assignment in addition to CAS target ranking (Ref 6:66).

Target correlation is the process of comparing new information about a reported unit to information about existing units to determine if the new information applies to an existing unit, or, if it implies the existence of a unit not previously detected. The battle assessment application of signature tables provides evaluation of a situation in terms of estimated combat time and force losses and is used to support the plan evaluation process. Ranking plans of action in terms of goals achieved and forces lost is the function of the plan evaluation process while the assignment of CAS aircraft to particular forward air controllers (FAC) is handled through the sortie assignment process (Ref 6:66).

Signature table techniques are applied to all of the above processes in the general manner described in Chapter II. An example of this application of the CAS target ranking process area is the set of FAC Target Importance Tables.

FAC Target Importance Tables

One purpose of a FAC is to direct air strikes against enemy targets. When the number of enemy targets exceeds the air strike resources available for scheduling by the FAC, a decision must be made as to which enemy target is most important. The FAC might rank order the targets and direct strikes on the targets beginning at the top of the list. However, since the battlefield is a dynamic arena, the relative importance of targets could very well change with time and a periodic update of the target list would be necessary. This process of rank ordering targets is one of the decision

processes modeled using signature table techniques.

Target Characterization

Each enemy target under consideration is characterized in terms of seven features:

1. Proximity to friendly unit engaged.
2. Target momentum.
3. Target speed.
4. Terrain cover.
5. Air defense security.
6. Aircraft effectiveness.
7. Target type.

A quantization scheme is defined for each of the features which maps the measured value, X_i, of the feature into one of the quantization levels available for that feature. The quantized values, F_i, are determined in accordance with the rules in Table I (Ref 4). Figure 4 illustrates the structure of this set of signature tables. The derived features from the three first level tables (T_1, T_2, T_3) are used as entries to tables T_4 and T_5. The output of signature table T_5 is the final evaluation of the target (target importance) and is used to rank order the available targets. Each of the signature tables T_1 through T_5 are displayed in Tables II through VI.

Example

The workings of the FAC target importance tables can now be demonstrated. First, assume a FAC must evaluate two targets to determine which target will be attacked by a particular CAS aircraft available to the FAC. A data set for each target is constructed and input to the table hierarchy for evaluation. Assume the data sets are as follows.

19
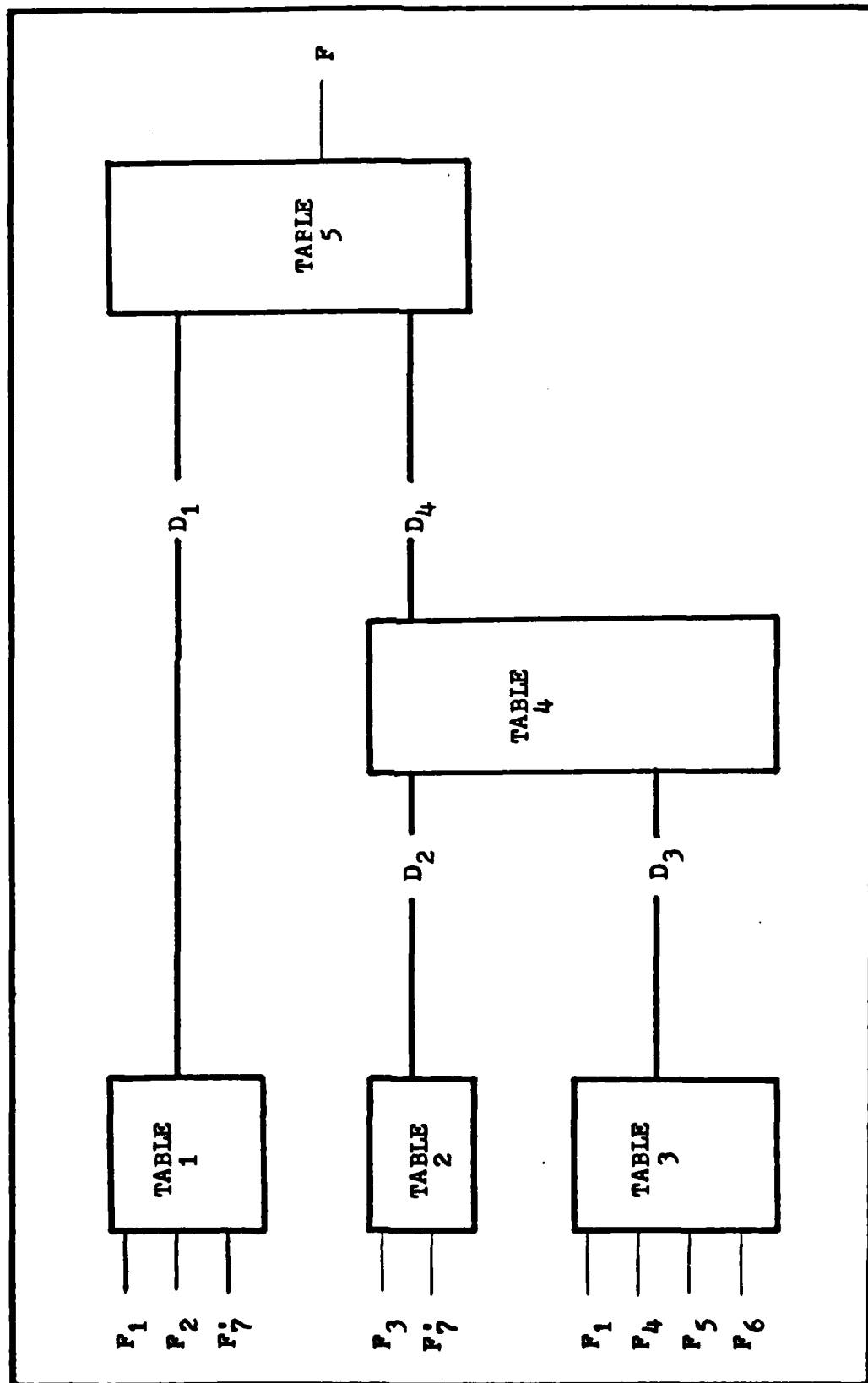
Fig 4. FAC Target Importance Tables Structure

TABLE I

Quantization Rules for Target Characterization

Measured Value, X_i	Quantized Value, F_i
$0 \leq X_1 \leq 3000$ (m)	1
$3000 < X_1 \leq 100,000$	2
$0 \leq X_2 \leq 10 \ (\frac{K_g - m}{s})$	1
$10 < X_2 \leq 100$	2
$0 \leq X_3 \leq 1$ (m/s)	1
$1 < X_3 \leq 1000$	2
X_4 = no terrain cover	1
X_4 = terrain cover	2
X_5 = no air defense	1
X_5 = air defense	2
$0 \leq X_6 \leq 11$ (%)	1
$11 < X_6 \leq 61$	2
$61 < X_6 \leq 100$	3
X_7 = Tank	1
X_7 = Armored Personnel Carrier	2
X_7 = Heavy Weapon	3
X_7 = Artillery	4

TABLE I Cont.

X_7 = Unknown	5
X_7 = Organic Air Defense	6
X_7 = Mobile Air Defense	7
X_7 = Fixed Air Defense	8

TABLE II
Signature Table 1 (Military Value Table)

F_1	F_2	F_8	D_1
1	1	1	2
1	1	2	1
1	1	3	1
1	1	4	3
1	1	5	1
1	1	6	1
1	1	7	2
1	1	8	4
1	2	1	1
1	2	2	1
1	2	3	1
1	2	4	2
1	2	5	1
1	2	6	1
1	2	7	1
1	2	8	4
2	1	1	4
2	1	2	4
2	1	3	3
2	1	4	5
2	1	5	2
2	1	6	2
2	1	7	5
2	1	8	4
2	2	1	5
2	2	2	4
2	2	3	5
2	2	4	4
2	2	5	3
2	2	6	3
2	2	7	3
2	2	8	4

TABLE III

Signature Table 2 (Target Mobility Table)

F_3	F_7	D_2
1	1	2
1	2	2
1	3	2
1	4	1
1	5	2
1	6	2
1	7	2
1	8	1
2	1	3
2	2	3
2	3	3
2	4	3
2	5	2
2	6	3
2	7	3
2	8	2

TABLE IV

Signature Table 3 (Target Vulnerability Table)

F_1	F_4	F_5	F_6	D_3
1	1	1	1	1
1	1	1	2	1
1	1	1	3	3
1	1	2	1	1
1	1	2	2	2
1	1	2	3	2
1	2	1	1	1
1	2	1	2	2
1	2	1	3	2
1	2	2	1	1
1	2	2	2	1
1	2	2	3	1
2	1	1	1	2
2	1	1	2	3
2	1	1	3	3
2	1	2	1	1
2	1	2	2	2
2	1	2	3	2
2	2	1	1	1
2	2	1	2	2
2	2	1	3	2
2	2	2	1	1
2	2	2	2	1
2	2	2	3	1

TABLE V

Signature Table 4 (Likelihood of Destruction Table)

D_3	D_2	D_4
1	1	2
1	2	2
1	3	1
2	1	3
2	2	2
2	3	2
3	1	3
3	2	2
3	3	2

TABLE VI

Signature Table 5 (Target Importance Table)

D_4	D_1	F
1	1	1
1	2	2
1	3	3
1	3	3
1	4	4
2	1	2
2	2	3
2	2	4
2	3	7
2	5	8
3	1	3
3	2	5
3	3	7
3	4	9
3	5	10

A = (4000,90,10, no terrain cover, no air defense
protection, 50, tank)

B = (1000,5, 0, no terrain cover, no air defense
protection, 75, APC)

The quantized data sets

A = (2, 2, 2, 1, 1, 2, 1)

B = (1, 1, 1, 2, 1, 3, 2)

are then entered in the first level tables. T_1 uses $(F_1$, F_2, $F_7)$ which for target A is (2, 2, 1). The D_1 corresponding to that signature is 5. For T_2, the signature is (2,1) for target A and D_2 = 3. The signature for T_3 is (2, 1, 1, 2) and D_3 = 3. The T_4 signature type is (D_3, D_2) which, in the case of target A is (3,3) and D_4 = 2. In the final table T_5, the entry signature is (2,5) and the target importance, F, is 8. Using the same procedure on target B, F = 2. Target A, there-fore, would be the target assigned to the available CAS aircraft.

Error

This process of using the signature tables to evaluate the available targets which results in a rank order list of targets is an attempt to explicitly model the DM's decision process in terms of quantifiable attributes. It is, in effect, a rough approximation of the value that a FAC places on a target characterized by the given attributes.

28
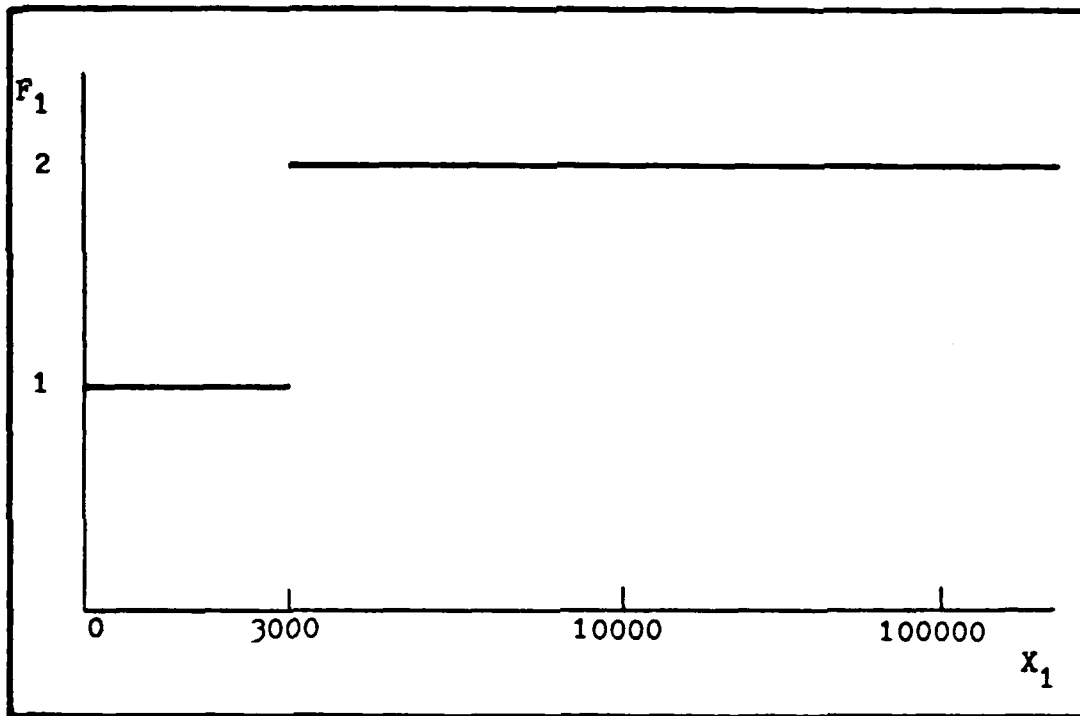
Fig 5. Proximity Quantization Step Function
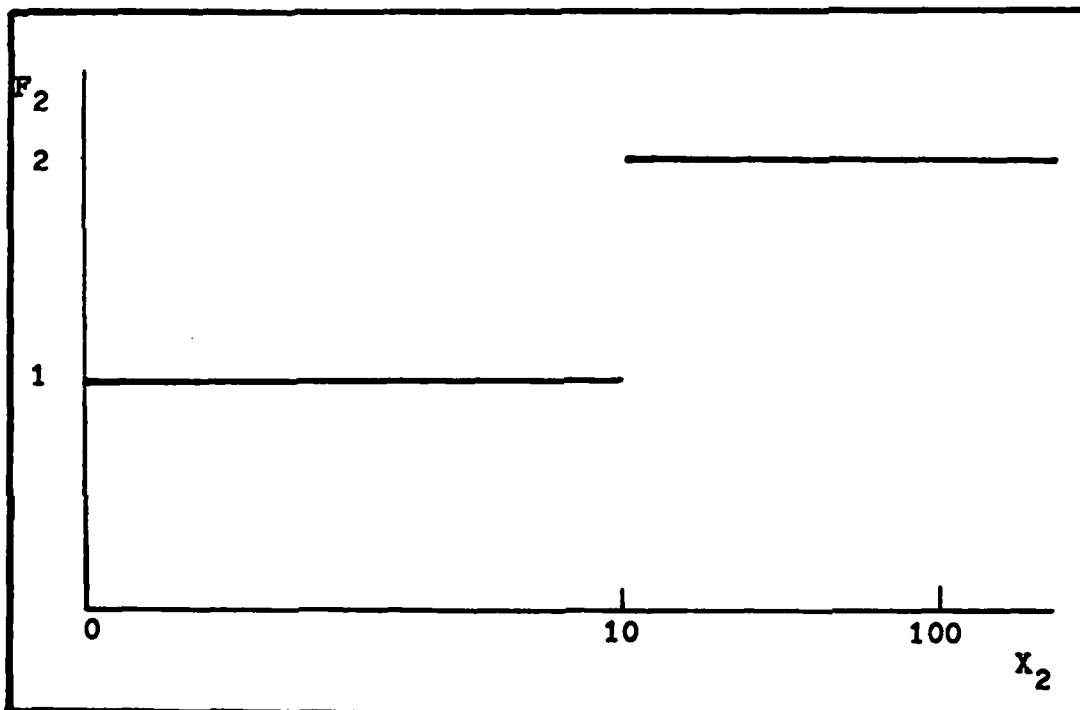
Fig 6. Momentum Quantization Step Function

The next subject is the question of error. Specifically, it is the question of the error that is generated by using a step function to quantize those attributes or features which are continuous.

For example, F_1 and F_2 are represented by the step functions in Figures 5 and 6. The proximity and momentum are not discrete, but are continuous variables. As such, what kind of error is generated by the attempt to represent these features by the given step function? Treatment of this question will be the subject of the next chapter.

IV. Methodology

The objective of this chapter is to show the development
of the error measures to be used in Chapter V. The error
to be measured is that which is generated by quantizing the
input features for a signature table evaluation instead of
using the continuous representation of the measured feature
and a value function evaluation. This error will be described
using three measures developed in this chapter. Two of
these measures are mean squared error (MSE) and percent
error (PE). The third will be a comparison of sets of target
rank order lists generated by the two evaluation schemes.

The starting point of the error measure development is
the derivation of a value function, U, from data in a particular
signature table. The MSE under two different sets of
assumptions concerning U is computed and then converted
into a "noise-to-signal" percent error. Finally, a simple
simulation model will be used to generate the sets of target
rank order lists for comparison. Further discussion of each
of these steps follows.

MSE Determination

The military value table, Table II, will be the subject
of the error measurement. The signature type evaluated by
this table consists of features F_1, F_2, and F_7. The quantization
rules that convert the measured data X_1, X_2, and X_7 into the
quantized values are listed in Table I. Some initial
assumptions about these measured data and their quantized
values need to be made at this point.

31

Assume there exists some function G_1 which transforms X_1 into the variable H_1 on the range $0 \le H_1 \le 2$ and, G_2 which transforms X_2 into the variable H_2 on the range $0 \le H_2 \le 2$. Further, assume these transformation are such that the quantization rules for X_1 and X_2 can be restated as follows:

$$0 \le H_1 \le 1, \; F_1 = 1 \tag{9}$$

$$1 < H_1 \le 2, \; F_1 = 2 \tag{10}$$

$$0 \le H_2 \le 1, \; F_2 = 1 \tag{11}$$

$$1 < H_2 \le 2, \; F_2 = 2 \tag{12}$$

Next, one of two assumptions about the table output, D_1, is necessary. The first possible assumption is that D_1 is the value function evaluated at the quantized values (F_1, F_2, F_7). That is,

$$D_1(F_1, F_2, F_7) = U(F_1, F_2, F_7) \tag{13}$$

The other possible assumption is that D_1 is the value function evaluated at some point (h_1, h_2, F_7) that lies within the specified quantization level:

$$D_1(F_1, F_2, F_7) = U(h_1, h_2, F_7) \tag{14}$$

Note here that X_7 is a discrete variable and cannot be represented by a continuous variable, hence it will be carried through the analysis with its quantized value F_7.

For example, consider a single feature example where the signature type consists of a feature with transformed value H. Then under the first assumption, the table output, D, at some quantized feature value F would equal U(F). If, however, the second assumption were true, D would equal U(h) where h was some value in the domain of H to which the quantization rules assigned the value F as shown in Figure 7.

One change was made to the original military value table before the value function U was derived. More specifically, a change was made to the quantization rules applied to X_7. It was discovered that by rearranging the values assigned to X_7, the table output becomes monotonically increasing for two combinations of F_1 and F_2 and minimized the positive-negative slope changes in the other two combinations. The quantization rules were changed to the system shown in Table VII, where F_7 and F_7' are the old and new quantized values, respectively. The resulting reordered military value table is illustrated in Table VIII.

The data from this new table was used to generate the value function U which was assumed to have the general form:

$$U(x,y,z) = Ax + By + Cz + Dxy + Exz + Fyz + G \qquad (15)$$

where

$x = F_1$

$y = F_2$

$z = F_7$

The method of least squares was used for this derivation and the results are presented in Chapter V.
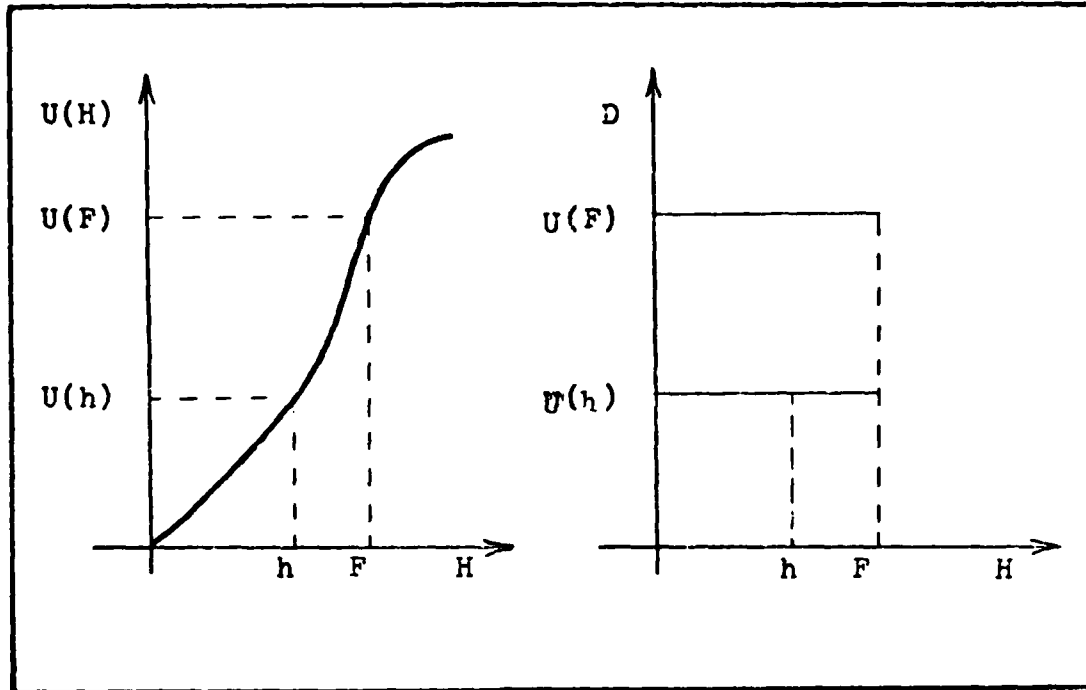
Fig 7. Single Feature Example

TABLE VII.

Revised Quantization Rules

X_7	F_7	F_7'
UNK	5	1
HVWP	3	2
ARTY	4	3
MADA	7	4
TANK	1	5
APC	2	6
OADA	6	7
FADA	8	8

TABLE VIII.

Revised Signature Table 1

F_1	F_2	F_7'	D_1
1	1	1	1
1	1	2	1
1	1	3	1
1	1	4	1
1	1	5	2
1	1	6	2
1	1	7	3
1	1	8	4
1	2	1	1
1	2	2	1
1	2	3	1
1	2	4	1
1	2	5	1
1	2	6	1
1	2	7	2
1	2	8	4
2	1	1	2
2	1	2	2
2	1	3	4
2	1	4	3
2	1	5	4
2	1	6	5
2	1	7	5
2	1	8	4
2	2	1	3
2	2	2	3
2	2	3	4
2	2	4	5
2	2	5	5
2	2	6	3
2	2	7	4
2	2	8	4

Once the function U was determined, error measurement could be made under each of the stated assumptions about the table output. For ease of reference, the case where $D_1 = U(F_1, F_2, F_7')$ will be Case 1 and where $D_1 = U(h_1, h_2, F_7')$ will be Case 2.

In Case 1, $U(F_1, F_2, F_7')$ should equal the table values D_1. However, since the method of least squares, in this case, does not yield a perfect fit, it will be assumed that U is exact and a new table constructed to match U evaluated at the quantized values of F_1, F_2, and F_7'. For Case 2, since D_1 does not equal $U(F_1, F_2, F_7')$, the least squares fit using the table data represents an approximation to U which includes error from the imperfect fit and from using F_1, F_2, and F_7' values instead of h_1, h_2, and F_7' as input data. For the purposes of this study, it will be assumed under Case 2 that the derived function U is the true value function and the table values D_1 are the results of evaluating U at the point (h_1, h_2, F_7').

In either case, the parameter of interest is MSE, which is defined as the expected value of the error squared where error is the difference between $U(H_1, H_2, F_7')$ and $D_1(F_1, F_2, F_7')$ (Ref 7:201).
Expressed mathematically:

$$MSE = E[\,|U(H_1, H_2, F_7') - D_1(F_1, F_2, F_7')|^2\,] \qquad (16)$$

If H_1, H_2, and F_7' are independent random variables with probability density function as follows,

$$f_1(H_1) = \begin{cases} \frac{1}{2}, & 0 \leq H_1 \leq 2 \\ 0, & \text{elsewhere} \end{cases} \tag{17}$$

$$f_2(H_2) = \begin{cases} \frac{1}{2}, & 0 \leq H_2 \leq 2 \\ 0, & \text{elsewhere} \end{cases} \tag{18}$$

$$p(F_7{}') = \begin{cases} \frac{1}{8}, & F_7{}' = 1,2,3,4,5,6,7,8 \\ 0, & \text{elsewhere} \end{cases} \tag{19}$$

then MSE can be expressed as

$$MSE = \sum_{F_7{}'=1}^{8} \int_0^2 \int_0^2 [U(H_1,H_2,F_7{}') - D_1(F_1,F_2,F_7{}')]^2 f_1(H_1) f_2(H_2) dH_1 dH_2$$

$$MSE = \frac{1}{32} \sum_{F_7{}'=1}^{F_7{}'=8} \int_0^2 \int_0^2 [U(H_1,H_2,F_7{}') - D_1(F_1,F_2,F_7{}')]^2 dH_1 dH_2 \tag{20}$$

This expression must be broken down as follows for evaluation

$$MSE = \frac{1}{32} \sum_{F_7{}'=1}^{8} \int_0^1 \int_0^1 [U(H_1,H_2,F_7{}') - D_1(1,1,F_7{}')]^2 dH_1 dH_2$$

$$+ \frac{1}{32} \sum_{F_7{}'=1}^{8} \int_1^2 \int_0^1 [U(H_1,H_2,F_7{}') - D_1(2,1,F_7{}')]^2 dH_1 dH_2$$

$$+ \frac{1}{32} \sum_{F_7'=1}^{8} \int_1^2 \int_0^1 [U(H_1,H_2,F_7') - D_1(1,2,F_7')]^2 dH_1 dH_2$$

$$+ \frac{1}{32} \sum_{F_7'=1}^{8} \int_0^2 \int_0^2 [U(H_1,H_2,F_7') - D_1(2,2,F_7')]^2 dH_1 dH_2 \quad (21)$$

Substituting the assumed general form for $U(H_1,H_2,F_7')$ and performing the integration and summation over the four regions yields:

$$\frac{128A^2}{3} + 64AB + 288AC + 288BC + \frac{256}{3}AD + 384AE + 288AF + 64AG +$$

$$\frac{128B^2}{3} + \frac{256BD}{3} + 288BE + 384BE + 384FB + 64BG + 816C^2 + 288CD +$$

$$1632CE + 1632CF + 288CG + 1088F^2 + \frac{512D^2}{9} + 384DE + 384DF +$$

$$64DG + 1088E^2 + 1632EF + 288EG + 288FG + 32G^2$$

$$- \sum_{F_7'=1}^{8} D_1(1,1,F_7') \ [A + B + 2CF_7'D + \frac{D}{2} + EF_7' + FF_7' + 2G - D_1(1,1,F_7')]$$

$$- \sum_{F_7'=1}^{8} D_1(2,1,F_7') \ [3A+B+2CF_7' + \frac{3D}{2} + 3EF_7' + FF_7' + 2G - D_1(2,1,F_7')]$$

$$- \sum_{F_7'=1}^{8} D_1(1,2,F_7') \ [A+3B+2CF_7' + \frac{3D}{2} + EF_7' + 3FF_7' + 2G - D_1(1,2,F_7')]$$

$$- \sum_{F_7'=1}^{8} D_1(2,2,F_7') [3A+3B+2CF_7' + \frac{9D}{2} + 3EF_7' + 3FF_7' + 2G - D_1(2,2,F_7')]$$

$$(22)$$

The numerical results under Case 1 and Case 2 are different and are contained in Chapter V.

Percent Error

Since MSE, as applied to this problem, is simply a number with no real meaning other than smaller is better, some way of attaching meaning was required. This was accomplished in two ways. The first was a percentage error measure which was defined as

$$PE = \frac{MSE}{MSS} \times 100 \qquad (23)$$

where MSS is the mean squared signal. The MSS under Case 1 and Case 2 assumptions are equal and can be expressed as

$$MSS = \frac{1}{32} \sum_{F_7' =1}^{8} \int_0^2 \int_0^2 U^2(H_1, H_2, F_7') dH_1 dH_2 \qquad (24)$$

After substituting the assummed general form of $U(H_1, H_2, F_7')$, MSS is as follows

$$MSS = \frac{128A^2}{3} + 64AB + 288AC + \frac{256AD}{3} + 384AE + 288AF$$

$$+ 64AG + \frac{128B^2}{3} + 288BE + \frac{256BD}{3} + 288BE$$

$$+ 384BF + 64BG + 816C^2 + 288CD + 1632CE + 1632CF$$

$$+ 288CG + \frac{513D^2}{9} + 384DE + 384DF + 64DG$$

$$+ 1088E^2 + 1632EF + 288EG + 1088F^2 + 288FG + 32G^2 \qquad (25)$$

39

PE, then, will give a measure of the error expressed as a percentage of the assumed evaluation function under each case of assumptions.

Evaluation Simulation

The second way of attaching meaning to the results is to generate some data sets and evaluate them using each method under each assumption case. In order to do that, however, a second evaluation function had to be determined. Since the output of table 1 is used as an input to signature table 5, a continuous output from signature table 1 would either have to be quantized for use by table 5 or signature table 5 needed to be treated like table 1. The latter method was chosen since re-quantizing the output from table 1 to match the allowable input for signature table 5 would destroy the error information of concern. Signature table 5 allows five quantization levels from table 1 which means the output of $U(H_1,H_2,F_7')$ would be quantized back into five levels and any error increase (or decrease) from using $U(H_1,H_2, F_7')$ instead of the signature table would get canceled. A general form for $U_1(D4, U(H_1,H_2,F_7'))$ was assumed to be Ax+By+C and a least squares fit obtained. Using the same procedures as for signature table 1, MSE and PE were computed and are contained in Chapter V.

An evaluation program was written to generate target data sets, evaluate them using both methods and assumption cases, and rank order the results of the evaluation and is contained in Appendix A. The data sets were generated through the use

of a uniform random variate generator and scaled to the appropriate range. For the signature table evaluation, each data set was quantized according to the quantization rules in Table I. For the function evaluation, the first two features were kept in their continuous form while the remaining five were quantized. The structure of the function evaluation is presented in Figure 8.

One final measure was included in the evaluation program. The Spearman rank correlation coefficient was computed and used as a measure of association between the rank ordering using the two methods. Further discussion of its use is presented in Chapter V.

In summary, a theoretical error measure, MSE, was developed. This measure was then converted into a percentage error measure, PE, which yields an error measure which is intuitively more meaningful. To reinforce that meaning, a sample of 30 targets is evaluated and rank ordered for comparison by each of the two evaluation schemes. Finally, a measure of association is used to determine if, in fact, there exists a significant difference in the rank ordering which results from each scheme.
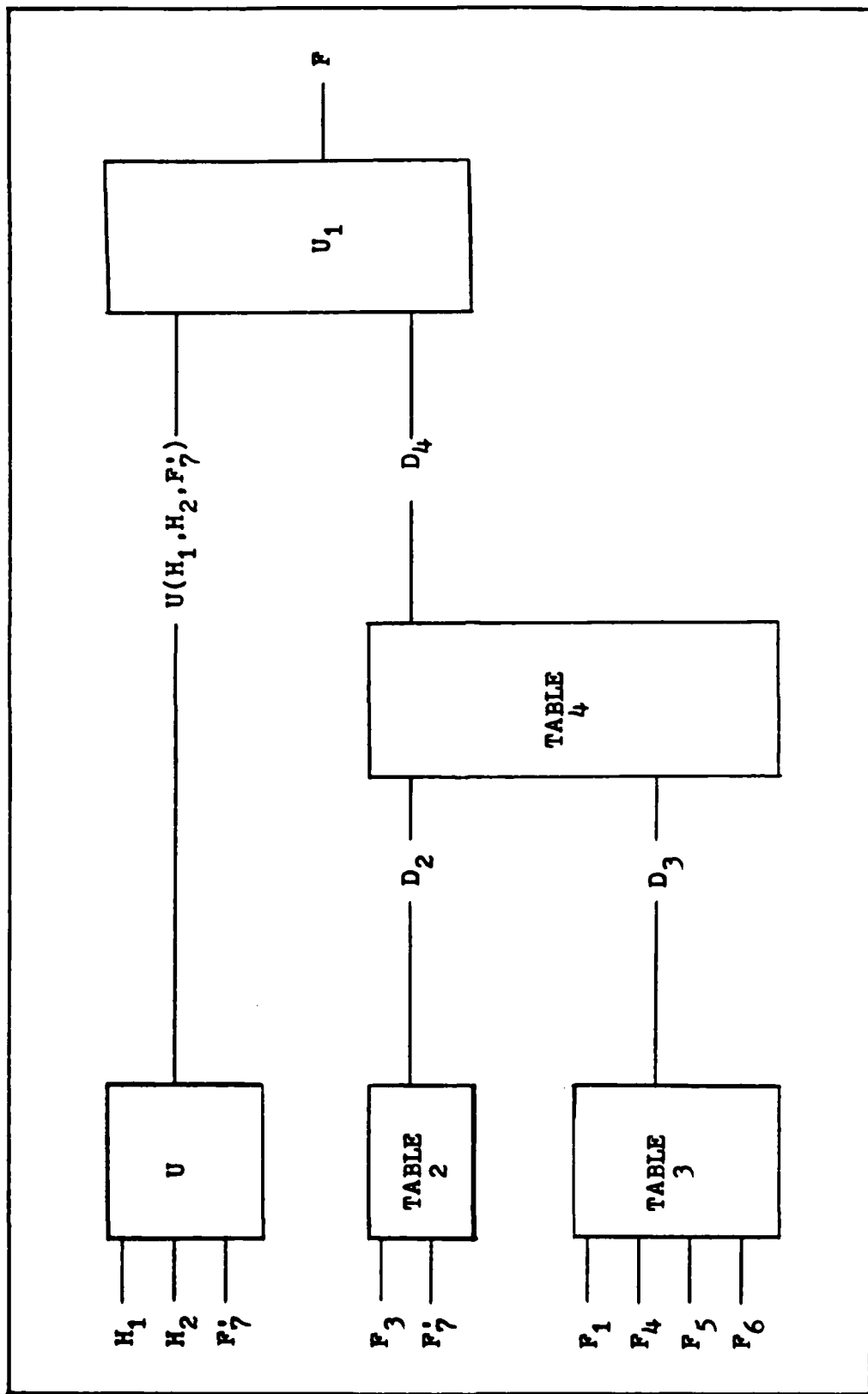
Fig 8. Function Evaluation Structure

42

V. Results

In this section, first the MSE for signature table 1 and
5 will be presented under Case 1 assumptions. This will
be followed by MSE for both tables under Case 2 assumptions.
Next, the computation of the percent error for both cases
followed by the evaluation simulation results. The last
item that will be covered is Spearman's rank order correlation
coefficient.

MSE

Under the Case 1 assumptions, the table output of a
given table represents some value function evaluated at the
quantized values. The least squares fit of data in signature
table 1 to the general form of equation (15) yielded
the following coefficients.

$$A = 1.6339286$$
$$B = -.1160714$$
$$C = .7708333$$
$$D = .6250000$$
$$E = -.1130952$$
$$F = -.1964286$$
$$G = -2.3437500$$

Using the resulting equation (15) a few signature table was
generated and is presented in Table IX. The effect of this
substitution is that $U(H_1, H_2, F_7')$ is now an exact fit to
the new military value table and quantization error under
these conditions can now be determined. The same procedure

TABLE IX

Signature Table 1 Revised for Case 1

F_1	F_2	F_7'	D1
1	1	1	.2604
1	1	2	.7217
1	1	3	1.1830
1	1	4	1.6443
1	1	5	2.1057
1	1	6	2.5670
1	1	7	3.0283
1	1	8	3.4896
1	2	1	.5729
1	2	2	.8378
1	2	3	1.1027
1	2	4	1.3676
1	2	5	1.6324
1	2	6	1.8973
1	2	7	2.1622
1	2	8	2.4271
2	1	1	2.4063
2	1	2	2.7545
2	1	3	3.1027
2	1	4	3.4509
2	1	5	3.7991
2	1	6	4.1473
2	1	7	4.4955
2	1	8	4.8438
2	2	1	3.3438
2	2	2	3.4955
2	2	3	3.6473
2	2	4	3.7991
2	2	5	3.9509
2	2	6	4.1027
2	2	7	4.2545
2	2	8	4.4063

was applied to signature table 5 for the reasons given in Chapter IV. The assumed form of the value function was $Ax + By + C$ where $X = D_4$ and $y = u(H_1, H_2, F_7')$. The coefficients generated by a least squares fit were

$A = 1.9$

$B = 1.4333$

$C = -3.1$

The revised signature table 5 is presented in Table X. Equation (14), using the adjusted signature table 1 values reduces to

$$MSE = \frac{1}{32} (186.6841 - 148.4483)$$

$$MSE = 1.1949$$

The MSE for signature table 5 is computed in the same way and yields

$$MSE = \frac{1}{15} (375.4919 - 365.0954)$$

$$MSE = 0.6931$$

For Case 2 assumptions, the derived value function is again assumed to be correct and the signature table values are the result of evaluating the function at points other than the quantized feature values. The mean squared error under Case 2 for signature table 1 is

$$MSE = \frac{1}{32} (186.6841 - 134.7680)$$

$$MSE = 1.6224$$

TABLE X.

Signature Table 5 Revised for Case 1

15	15	15
D_4	D_1	F
1	1	.2333
1	2	1.6666
1	3	3.0999
1	4	4.5332
1	5	5.9665
2	1	2.1333
2	2	3.5666
2	3	4.9999
2	4	6.4332
2	5	7.8665
3	1	4.0333
3	2	5.4666
3	3	6.8999
3	4	8.3332
3	5	9.7665

46

Once again, the same computation for the original signature table 5 yields

$$MSE = \frac{1}{15} \quad (375.4919 - 360.9513)$$

$$MSE = .9694$$

It can be seen that under Case 1 assumptions, where the table output was, in fact, the value function evaluated at the quantized features, the MSE was better (smaller) than when the table output was determined as in Case 2.

Percent Error

Recall from Chapter IV, equation (23) which defined a percent error measure. This measure expresses the quantization error, MSE, as a percent of the mean squared signal, MSS (equation 24). Under Case 1, MSS and PE for signature table 1 are

$$MSS = \frac{1}{32} \quad (186.6841)$$

$$MSS = 5.8339$$

$$PE = \frac{1.1949}{5.8339} \times 100$$

$$PE = 20.48\%$$

The MSS and PE for signature table 5 are

$$MSS = 25.0328$$

$$PE = 2.77\%$$

Applying Case 2 assumptions, the results changed as
follows. For signature table 1

MSS = 5.8339

PE = 27.81%

Calculations for signature table 5 yield

MSS = 25.0328

PE = 3.87%

These results indicate that in either assumption
case, the initial feature quantization error is at least 20% (PE
for table 1 under Case 1). Signature table 5 only adds
another 3-4% to the initial quantization error. The
evaluation simulation was written in an attempt to give
a better feel for what the results thus far mean.

Evaluation Simulation

Under Case 1 assumptions, the evaluations and rank order
listings of the 30 sample targets are presented in Table XI
for both the signature table and value function methods.
Table XII contains the same information for Case 2 assumptions.
A significant result in the signature table ranking is the
number of ties that resulted under Case 2 assumptions.
Out of 30 targets evaluated, a total of 28 involved ties.
This indicates that this method of evaluation must include
a means of settling ties. A tie between the evaluation of
two or more targets defeats the purpose of the evaluation

TABLE XI

Case 1 Rank Order List

Target	Signature Table		Value Function	
Number	F	RANK	U	RANK
1	5.9	24.0	2.8620	14.0
2	1.5	5.0	-.0845	6.0
3	4.2	17.0	4.3709	24.0
4	.5	1.0	-.6225	5.0
5	6.1	26.0	4.4993	26.0
6	2.7	7.0	2.0878	10.0
7	5.6	22.5	3.7377	22.0
8	3.4	12.0	3.5396	20.0
9	5.1	19.0	4.5613	27.0
10	6.8	29.0	5.4545	29.0
11	6.1	26.0	4.1362	23.0
12	3.0	9.0	2.9885	16.0
13	1.1	2.5	-.7897	3.0
14	4.2	18.0	3.3920	19.0
15	5.5	21.0	2.7129	12.0
16	3.8	14.0	2.7735	13.0
17	3.0	9.0	1.9911	9.0
18	6.1	26.0	3.3367	18.0
19	8.7	30.0	7.4672	30.0
20	4.0	15.0	.5291	7.0
21	6.4	28.0	5.0821	28.0
22	4.1	16.0	2.9863	15.0
23	1.5	5.0	-1.6174	1.0
24	1.1	2.5	-.7503	4.0
25	5.6	22.5	3.2957	17.0
26	5.2	20.0	4.4379	25.0
27	3.8	13.0	3.5408	21.0
28	3.0	9.0	1.7593	8.0
29	1.5	5.0	-1.3775	2.0
30	3.1	11.0	2.6527	11.0

TABLE XII

Case 2 Rank Order List

Target	Signature Table		Value Function	
Number	F	RANK	U	RANK
1	7.0	23.5	2.8620	14.0
2	2.0	8.0	-.0845	6.0
3	7.0	23.5	4.3709	24.0
4	1.0	2.0	-.6225	5.0
5	8.0	27.5	4.4993	26.0
6	2.0	8.0	2.0878	10.0
7	5.0	19.5	3.7377	22.0
8	2.0	8.0	3.5396	20.0
9	7.0	23.5	4.5613	27.0
10	7.0	23.5	5.4545	29.0
11	8.0	27.5	4.1362	23.0
12	2.0	8.0	2.9885	16.0
13	1.0	2.0	-.7897	3.0
14	4.0	15.5	3.3920	19.0
15	5.0	19.5	2.7129	12.0
16	3.0	13.0	2.7735	13.0
17	2.0	8.0	1.9911	9.0
18	8.0	27.5	3.3367	18.0
19	9.0	30.0	7.4672	30.0
20	4.0	15.5	.5291	7.0
21	8.0	27.5	5.0821	28.0
22	4.0	15.5	2.9863	15.0
23	2.0	8.0	-1.6174	1.0
24	1.0	2.0	-.7503	4.0
25	5.0	19.5	3.2957	17.0
26	5.0	19.5	4.4379	25.0
27	4.0	15.5	3.5409	21.0
28	2.0	8.0	1.7593	8.0
29	2.0	8.0	-1.3775	2.0
30	2.0	8.0	2.6527	11.0

if the only targets to be ranked are tied. At best, the effectiveness of the system is drastically impaired assumming not all the targets to be ranked tie in their evaluations. The purpose of the evaluation is to prioritize the allocation of weapon resources and if the evaluation method used yields a tie, the decision maker is left at square one.

This apparent lack of fidelity in the evaluation from the signature table method under Case 2 assumptions is not a problem in the value function method and is much less significant in the signature table method under Case 1 assumptions. The value function method yielded no ties while the Case 1 signature table method yielded only 13 of the 30 targets involved in ties. In both cases, the fidelity of the evaluation systems is improved over the Case 2 signature table method.

The data sets that were evaluated in the previous runs were generated by assuming each of the seven features were uniformly distributed across the appropriate ranges. In a further attempt to visualize the results of quantizing the input variables, two more sets of data were generated. In one set, features X_1 and X_2 were drawn from an exponential distribution with mean = 1, while the other features were generated as before. The computer program modification is contained in Appendix B. The results are contained in Tables XIII and XIV for Case 1 and Case 2 assumptions, respectively. The signature table method under Case 1 resulted in 16 of 30 targets involved in ties. Under Case 2, 24 of the 30 targets were involved in ties.

51

TABLE XIII

Case 1 Rank Order List (Exponential)

Target	Signature Table		Value Function	
Number	F	RANK	U	RANK
1	0.5	2.5	-.2609	5.0
2	4.1	21.0	1.2579	9.0
3	7.6	29.0	6.9179	28.0
4	4.0	19.0	2.3891	14.0
5	3.1	12.5	2.1499	13.0
6	3.1	12.5	3.2631	20.0
7	3.1	12.5	1.8421	10.0
8	4.4	24.0	4.8327	26.0
9	2.3	8.0	-.1612	7.0
10	3.1	16.0	2.9027	17.0
11	3.1	12.5	2.0188	11.0
12	3.7	17.5	3.9516	23.0
13	6.1	26.0	5.9806	27.0
14	1.1	5.0	-1.4396	3.0
15	1.1	4.0	-1.0317	4.0
16	-.2	1.0	-2.2308	1.0
17	6.1	26.0	4.6451	24.0
18	3.1	12.5	2.4397	15.0
19	6.9	28.0	7.0586	29.0
20	.5	2.5	-.2458	6.0
21	3.7	17.5	3.0716	18.0
22	1.5	6.0	-1.5888	2.0
23	4.1	21.0	3.2105	19.0
24	4.2	23.0	3.5385	21.0
25	2.7	9.0	2.0788	12.0
26	1.9	7.0	.7075	8.0
27	9.0	30.0	8.9918	30.0
28	6.1	26.0	4.6915	25.0
29	4.1	21.0	3.6414	22.0
30	3.1	12.5	2.8702	16.0

TABLE XIV

Case 2 Rank Order List (Exponential)

Target NUMBER	Signature Table F	RANK	Value Function U	RANK
1	1.0	2.5	-.2609	5.0
2	4.0	20.5	1.2579	9.0
3	7.0	27.0	6.9179	28.0
4	4.0	20.5	2.3891	14.0
5	2.0	10.0	2.1499	13.0
6	2.0	10.0	3.2631	20.0
7	2.0	10.0	1.8421	10.0
8	4.0	20.5	4.8327	26.0
9	2.0	10.0	-.1612	7.0
10	3.0	16.0	2.9027	17.0
11	2.0	10.0	2.0188	11.0
12	4.0	20.5	3.9516	23.0
13	7.0	27.0	5.9806	27.0
14	1.0	2.5	-1.4396	3.0
15	2.0	10.0	-1.0317	4.0
16	1.0	2.5	-2.2308	1.0
17	7.0	27.0	4.6451	24.0
18	2.0	10.0	2.4397	15.0
19	7.0	27.0	7.0586	29.0
20	1.0	2.5	-.2458	6.0
21	4.0	20.5	3.0716	18.0
22	2.0	10.0	-1.5888	2.0
23	4.0	20.5	3.2105	19.0
24	4.0	20.5	3.5385	21.0
25	2.0	10.0	2.0788	12.0
26	2.0	10.0	.7075	8.0
27	10.0	30.0	8.9918	30.0
28	7.0	27.0	4.6915	25.0
29	4.0	20.5	3.6414	22.0
30	2.0	10.0	2.8702	16.0

The second additional set of data was generated by taking X_1 and X_2 from a normal distribution with mean of 0.75 and standard deviation 0.25. The required program modification is contained in Appendix C. The results of these runs are contained in Tables XV and XVI. Using the signature table method under Case 1 assumptions, 17 of the 30 targets evaluated were tied. Under Case 2 assumptions, 29 of the 30 targets resulted in ties. The value function method in both additional data sets, of course, yielded no ties.

Once again, the lack of fidelity in the signature table method under either set of assumptions becomes apparent. In all three data sets, the signature table method under Case 2 assumptions results in a significant number of ties. These ties, as stated before, provide no real solution to the DM's rank ordering problem.

Spearman Rank Correlation Coefficient

The last measure of merit to be applied to these methods of target evaluation and ranking comes from the field of non-parametric statistics. This measure, Spearman's rank correlation coefficient, can be used to test the degree of association between series of rank ordered data. The rank correlation coefficient, r_s, is computed using the following equation

$$r_s = 1 - \frac{6 \sum_{i=1}^{n} d_i^2}{n(n^2-1)} \tag{26}$$

where d_i is the difference in the ranks of the same target by the two methods being compared (Ref 2:550). The rank

54

TABLE XV

Case 1 Rank Order List (Normal)

Target	Signature Table		Value Function	
Number	F	RANK	U	RANK
1	1.1	2.5	-1.0041	2.0
2	2.7	17.0	1.3397	11.0
3	4.4	21.5	4.4936	25.0
4	2.5	16.0	2.2792	18.0
5	4.4	21.5	4.3279	24.0
6	5.0	25.0	5.2359	26.0
7	5.6	26.0	3.6160	21.0
8	1.8	10.5	1.7026	13.0
9	4.4	21.5	3.2646	19.0
10	2.4	13.5	1.7290	15.0
11	8.9	30.0	8.0302	30.0
12	2.4	13.5	1.7036	14.0
13	2.4	13.5	2.1417	17.0
14	5.7	27.5	5.9257	27.0
15	1.2	4.0	-.0985	5.0
16	1.1	2.5	-1.2198	1.0
17	.5	1.0	.0579	6.0
18	2.4	13.5	1.1676	10.0
19	1.5	5.0	.9480	9.0
20	1.7	8.0	-.1529	4.0
21	6.9	29.0	6.0758	28.0
22	1.7	8.0	-.3551	3.0
23	1.8	10.5	.5340	7.0
24	4.4	21.5	3.6504	22.0
25	5.7	27.5	6.4688	29.0
26	3.7	19.0	3.3972	20.0
27	1.7	8.0	.6912	8.0
28	1.6	6.0	1.3466	12.0
29	3.4	18.0	4.1435	23.0
30	4.6	24.0	1.8486	16.0

TABLE XVI

Case 2 Rank Order List (Normal)

Target	Signature Table		Value Function	
Number	F	RANK	J	RANK
1	2.0	11.0	-1.0041	2.0
2	2.0	11.0	1.3397	11.0
3	4.0	21.5	4.4936	25.0
4	2.0	11.0	2.2792	18.0
5	4.0	21.5	4.3279	24.0
6	5.0	25.5	5.2359	26.0
7	5.0	25.5	3.6160	21.0
8	2.0	11.0	1.7026	13.0
9	4.0	21.5	3.2646	19.0
10	2.0	11.0	1.7290	15.0
11	9.0	30.0	8.0302	30.0
12	2.0	11.0	1.7036	14.0
13	2.0	11.0	2.1417	17.0
14	7.0	28.0	5.9257	27.0
15	1.0	2.0	-.0985	5.0
16	2.0	11.0	-1.2198	1.0
17	1.0	2.0	.0579	6.0
18	2.0	11.0	1.1676	10.0
19	2.0	11.0	.9480	9.0
20	2.0	11.0	-.1529	4.0
21	7.0	28.0	6.0758	28.0
22	2.0	11.0	-.3551	3.0
23	2.0	11.0	.5340	7.0
24	4.0	21.5	3.6504	22.0
25	7.0	28.0	6.4688	29.0
26	4.0	21.5	3.3972	20.0
27	2.0	11.0	.6912	8.0
28	1.0	2.0	1.3456	12.0
29	2.0	11.0	4.1435	23.0
30	4.0	21.5	1.8486	16.0

correlation coefficient can be used in this situation to determine if there exists a statistical difference between the evaluation methods under consideration. Specifically, the null hypothesis was no association between the two methods. Acceptance of the null hypothesis would imply a significant difference existed in the valuation methods. The critical values of r_s for a sample size of 30 are 0.305 and 0.432 for a 0.95 and 0.99 confidence level (Ref 2:A46). The null hypothesis of no association would then be rejected if r_s was greater than or equal to the appropriate critical value. Under Case 1 assumptions, calculations yield r_s = .8396 which is well beyond the rejection region for both confidence levels. The value of r_s for Case 2 assumptions is .8254, also well into the rejection region. The null hypothesis of no association would be rejected under either set of assumptions at the 0.01 level of significance. This result means that the signature table method of evaluation does result in a rank order listing of targets that is associated to a statistically significant level with the continuous value function evaluation. A measure of that degree of association has already been presented as the MSE and subsequent expression as percent error.

The application of this same measure to the additional data sets yields similar results. These results are summarized in Table XVII.

TABLE XVII

Spearman Rank Correlation Coefficients

Data Set	Case 1	Case 2
Uniform	0.8396	0.9254
Exponential	0.9211	0.8776
Normal	0.9218	0.8249

In all cases, the null hypothesis can be rejected and, there-fore,the signature table representation of the value function results in evaluations that are statistically similar.

VI. <u>Conclusions and Recommendations</u>

The signature table is a representation of a value function which has been evaluated at certain input levels. This is the major assumption of this study and if it is valid, significant benefits could be derived form replacing the signature table with the value function it represents.

First, the fidelity of the resultant figure would be greatly improved. For instance, using signature tables in the evaluation of targets as presented, yeilded a much higher probability of a tie between targets than did use of the assumed value function. Since the purpose of the scheme is to simulate a FAC rank ordering available targets for strikes, a ties does no good. The decision maker will not accept a tie in the evaluation of targets and if one results from an artificial evaluation method, a tie breaker must be included. The value function method needs no such tie breaking scheme. Second, there would be no need for quantization rules to be applied to the measured continuous features, X_n, in the value function evaluation scheme. For those features that are discrete to begin with, such as target type, a means of quantizing the measured value into data would still be necessary. One of the stated limitations of signature table performance is that it "is directly related to the discrete ranges of the input features" (Ref 6:64). This limitation could be negated through use of a value function. Third, there is the question of storage space required for each of the two methods. Both require an input data set array. However, the value function method

59

requires only storage of the equation coefficients and a function statement to perform the evaluation while the signature table method requires storage of the output features and a table search routine to find the required output. In the case of the evaluation simulation written for this study, the signature table method required storage of 96 table values and search routine of five lines. The value function method required storage of ten values (coefficients) and two function statements. The savings results from the number of values that must be stored. In this instance a storage savings of almost 90% resulted. This research has shown that there is sufficient merit to the application of a value function as a replacement for signature table evaluation to warrant further formal consideration. It has been shown, subject to the stated assumptions, that there is a quantization error that can be eliminated by replacing the signature evaluation scheme with a value function scheme. The benefits interms of improved evaluation fidelity and decreased storage requirements have been demonstrated. The value function method of evaluation clearly has merit and should be pursued.

Subject to the stated assumptions, these benefits and the percentage error figures previously reported lead to the conclusion that a value function replacement for signature table evaluation should be investigated in greater detail. Specifically, it is the recommendation of this study that a formal application of multiattribute utility assessment procedures such as those described by Keeney and Raiffa to a specific signature table evaluation scheme be

accomplished (Ref 1). This line of research could have significant impact on the artificial intelligence techniques currently in use to model the decision maker in computer simulations.

The concept of a decision maker being risk averse, neutral, or prone could be incorporated into the simulation using utility functions in place of signature tables and the results analyzed. This is not to imply that treatment of risk is absent in the signature table method; however, it would be more easily seen and interpreted using utility functions.

Bibliography

1. Keeney, Ralph L. and Howard Raiffa. <u>Decisions With Multiple Objectives: Preferences and Value Tradeoffs</u>. New York: John Wiley and Sons, 1976.

2. Mendenhall, William and Richard L. Scheaffer. <u>Mathematical Statistics with Applications</u>. North Scituate: Duxbury Press, 1973.

3. Page, Carl V. "Heuristics for Signature Table Analysis as a Pattern Recognition Technique," <u>IEEE Transactions on Systems, Man, and Cybernetics</u>, <u>SMC-7(2)</u>: 77-86 (February 1977).

4. Reiss, Royce H., Tactical Systems Division, ACS/Studies and Analysis (personal correspondence). Pentagon, Washington D.C., 29 August, 1979.

5. Samuel, A. L. "Some Studies in Machine Learning Using the Game of Checkers. II-Recent Progress," <u>IBM Journal of Research and Development</u>, <u>6</u>: 601-617 (November 1967).

6. "TAC ASSESSOR Methodology Manual," prepared for Tactical Systems Division, ACS/SA, Headquarters USAF, by General Research Corporation, December 1978.

7. Taub, Herbert and Donald L. Schilling. <u>Principles of Communication Systems</u>. New York: McGraw-Hill Book Co., 1971.

Appendix A

```
      PROGRAM TGTIMP(INPUT,OUTPUT,TAPE8)
      INTEGER TABLE1,TABLE5
      INTEGER TABLE2,TABLE3,TABLE4
      INTEGER X,XD,FD
      DIMENSION XC(7), XD(7), TABLE1(2,2,8), TABLE2(2,8),
     CTABLE3(4,2,3), TABLE4(3,3), TABLE5(3,5), X(7),
     CRD1(30),RD2(30),RC(30),IR(30),R(30),RANKRC(30),RANKRD1(30),
     CRANKRD2(30),XD1(7)
      COMMON TABLE1,TABLE2,TABLE3,TABLE4,TABLE5
801   FORMAT (8I3)
802   FORMAT (3I3)
803   FORMAT (5I3)
901   FORMAT (1H1,13X,"    TARGET",7X,"SIGNATURE TABLE",9X,
     C"VALUE FUNCTION",/)
902   FORMAT (13X,"    NUMBER",7X,"F",10X,"RANK",7X,"U",11X,
     C"RANK",/,/)
903   FORMAT (18X,I2,8X,F4.1,9X,F4.1,5X,F8.1,7X,F4.1)
904   FORMAT (1H1)
905   FORMAT (8(T3,5X),I3,2(5X,F8.4),5(5X,I3))
      READ (8,801) (((TABLE1(I,J,K),K=1,8),J=1,2),I=1,2)
      READ (8,801) ((TABLE2(I,J),J=1,8),I=1,2)
      READ (8,802) (((TABLE3(I,J,K),K=1,3),J=1,2),I=1,4)
      READ (8,802) ((TABLE4(I,J),J=1,3),I=1,3)
      READ (8,803) ((TABLE5(I,J),J=1,5),I=1,3)
      WRITE 904
      DO 100 N=1,30
         CALL DATASET (XC,XD)
         WRITE 905,N,(XD(J),J=1,7),N,XC(1),XC(2),(XD(J),J=1,7)
         CALL TABD (XD,ID4,FD)
         CALL TABC (XC,ID4,FC)
         RD1(N)=FD
         RC(N)=FC
      DO 15 I=1,7
      XD1(I)=XD(I)
15    CONTINUE
         CALL TABC (XD1,ID4,FC)
         RD2(N)=FC
100   CONTINUE
      CALL NMRANK (RC,N,0.0,IR,R,RANKRC,S,T)
      CALL NMRANK (RD1,N,0.0,IR,R,RANKRD1,S,T)
      CALL NMRANK (RD2,N,0.0,IR,R,RANKRD2,S,T)
      WRITE 901
      WRITE 902
      DO 300 I=1,N
      WRITE 903,I,RD2(I),RANKRD2(I),RC(I),RANKRC(I)
300   CONTINUE
      WRITE 901
      WRITE 902
      DO 400 I=1,N
      WRITE 903,I,RD1(I),RANKRD1(I),RC(I),RANKRC(I)
400   CONTINUE
      D1=D2=0
```

```
      DO 40 I=1,N
          D1=D1+(RANKRC(I)-RANKRD1(I))**2
          D2=D2+(RANKRC(I)-RANKRD2(I))**2
40    CONTINUE
      R1=1-6*D1/(N*(N**2-1))
      R2=1-6*D2/(N*(N**2-1))
      PRINT*,1H1,"R1=   ",R1,"       R2=   ",R2
200   STOP "END OF TGTIMP"
      END
      SUBROUTINE DATASET (XC,XD)
      DIMENSION XC(7),XD(7)
      INTEGER XD
      DO 10 I=1,7
          A=RANF(D)
          IF (I.GT.5) GO TO 11
          XD(I)=2*A+1
          GO TO 13
11        IF (I.GT.6) GO TO 12
          XD(I)=3*A+1
          GO TO 13
12        XD(I)=8*A+1
13        IF (I.GT.2) GO TO 14
          XC(I)=2*A
          GO TO 10
14        XC(I)=XD(I)
10    CONTINUE
      RETURN
      END
      SUBROUTINE TABD (X,ID4,FD)
      COMMON TABLE1,TABLE2,TABLE3,TABLE4,TABLE5
      INTEGER TABLE1,TABLE5
      INTEGER TABLE2,TABLE3,TABLE4
      INTEGER X,FD
      DIMENSION X(7),TABLE1(2,2,8),TABLE2(2,8),TABLE3(4,2,3),
     CTABLE4(3,3),TABLE5(3,5)
      ID1=TABLE1(X(1),X(2),X(7))
      ID2=TABLE2(X(3),X(7))
      IF ((X(1).EQ.1).AND.(X(4).EQ.1)) GO TO 30
      GO TO 31
30    I=1
      GO TO 39
31    IF ((X(1).EQ.1).AND.(X(4).EQ.2)) GO TO 32
      GO TO 33
32    I=2
      GO TO 39
33    IF ((X(1).EQ.2).AND.(X(4).EQ.2)) GO TO 34
      I=3
      GO TO 39
34    I=4
39    ID3=TABLE3(I,X(5),X(6))
      ID4=TABLE4(ID3,ID2)
      FD=TABLE5(ID4,ID1)
      RETURN
```

64

```
END
SUBROUTINE TABC (X,ID4,FC)
DIMENSION X(7)
A=1.6339286
B=-.11507143
C=.7708333
D=.625
E=-.11309524
F=-.19642857
G=-2.34375
DA=A*X(1)+B*X(2)+C*X(7)
DB=D*X(1)*X(2)+E*X(1)*X(7)+F*X(2)*X(7)
D1=DA+DB+G
A=1.9
B=1.4333
C=-3.1
FC=A*ID4+B*D1+C
RETURN
```

```
      SUBROUTINE DATASET (XC,XD)
      DIMENSION XC(7),XD(7)
      INTEGER XD
      DO 10 I=1,7
         A=RANF(0)
         B=-ALOG(A)
         IF (I.GT.2) GO TO 11
         XD(I)=B+1
         IF (XD(I).GT.2) GO TO 18
         GO TO 12
18       XD(I)=2
         GO TO 12
11       IF (I.GT.5) GO TO 13
         XD(I)=2*A+1
         GO TO 12
13       IF (I.GT.6) GO TO 14
         XD(I)=3*A+1
         GO TO 12
14       XD(I)=8*A+1
12       IF (I.GT.2) GO TO 15
         XC(I)=B
         IF (XC(I).GT.2) GO TO 16
         GO TO 10
16       XC(I)=2
         GO TO 10
15       XC(I)=XD(I)
10    CONTINUE
      RETURN
      END
```

```
      SUBROUTINE DATAS-1 (XC,XD)
      DIMENSION XC(7),XD(7)
      INTEGER XD
      DO 10 I=1,7
          A=RANF(0)
          CALL NORMAL (.75,.25,X1)
          IF (I.GT.2) GO TO 11
          XD(I)=X1+1
          IF (XD(I).GT.2) GO TO 18
          IF (XD(I).LT.0) GO TO 19
          GO TO 12
18        XD(I)=2
          GO TO 12
19        XD(I)=0
          GO TO 12
11        IF (I.GT.5) GO TO 13
          XD(I)=2*A+1
          GO TO 12
13        IF (I.GT.6) GO TO 14
          XD(I)=3*A+1
          GO TO 12
14        XD(I)=8*A+1
12        IF (I.GT.2) GO TO 15
          XC(I)=X1
          IF (XC(I).GT.2) GO TO 16
          IF (XC(I).LT.0) GO TO 17
          GO TO 10
16        XC(I)=2
          GO TO 10
17        XC(I)=0
          GO TO 10
15        XC(I)=XD(I)
10    CONTINUE
      RETURN
      END


      SUBROUTINE NORMAL (EX,STD,X1)
10    C=RANF(0)
      V1=2*C-1
      D=RANF(0)
      V2=2*D-1
      S=V1**2+V2**2
      IF (S.GE.1) GO TO 10
      RN1=V1*SQRT((-2.0*ALOG(S))/S)
      X1=EX+PN1*STD
      RETURN
      END
```

VITA

David Scott Prahler was born on 14 June 1949 in Austin, Texas. He graduated from high school at Randolph Air Force Base, Texas in 1967 and attended the University of Texas at Austin for one year. He entered the United States Air Force Academy in June 1968 and graduated with a Bachelor of Science degree in Engineering Management and a commission in the USAF in June 1972. After receiving his initial weapons controller training at Tyndall AFB, Florida, he was assigned to the 23rd Air Division, Duluth IAP, Minnesota as a crew weapons controller. He was transferred to the 621st Tactical Control Squadron in December 1974 where he served as Control and Reporting Center Training Officer and Standardization/ Evaluation Controller. His next assignment was with the USAF Interceptor Weapons School at Tyndall AFB as a Squadron Weapons Instructor from December 1975 until he entered the School of Engineering, Air Force Institute of Technology, in August 1978.

Permanent address: 583 Kamoku St.
                              Apt 3207
                              Honolulu, Hawaii
                                      96825

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER AFIT/GST/EE/80-1	2. GOVT ACCESSION NO. AD-A085 711	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) QUANTIZATION ERROR DUE TO SIGNATURE TABLE EVALUATIONS		5. TYPE OF REPORT & PERIOD COVERED
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) DAVID S. PRAHLER CAPT USAF		8. CONTRACT OR GRANT NUMBER(s)
9. PERFORMING ORGANIZATION NAME AND ADDRESS Air Force Institute of Technology (AFIT-EN) Wright-Patterson AFB, OH 45433		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS AF/SAGR Tactical Systems Division Pentagon, Washington DC 20330		12. REPORT DATE
		13. NUMBER OF PAGES 68
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS. (of this report)
		15a. DECLASSIFICATION DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

Approved for public release; distribution limited IQW AFR 190-17

JOSEPH P. HIPPS, MAJOR, USAF
Director of Public Affairs

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Signature Tables
Value Function
Mean-Squared Error
Quantization Error
C^2 Simulation

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

The signature tables used in TAC ASSESSOR were assumed to be representations of some corresponding value functions evaluated at certain levels of the input features. The theoretical mean-squared-error due to this substitution was calculated for signature table 1 and table 5 in the FAC Target Importance Tables. The calculations were performed under two different assumption concerning the assumed value functions.

DD $_{1\ JAN\ 73}^{FORM}$ 1473   EDITION OF 1 NOV 65 IS OBSOLETE

Case 1 was that the signature table values resulted from evaluating the assumed value function at the given input feature values. Case 2 was that the function was evaluated at some feature values in the given quantization level indicated by the input feature. The mean squared error was then expressed as a percentage of the assumed mean squared signal. The percent errors derived were 20.48% and 27.81% for signature table 1 under Case 1 and 2 assumptions, respectively.

Next, an evaluation simulation was performed on three sets of 30 targets whose features were generated using uniform, exponential, and normal random variate generators scaled to the appropriate feature limits. The major result of this was that a significant number of the targets were involved in ties when evaluated using the signature table method. The number of ties ranged from 13 out of 30 up to 28 out of 30 targets involved in evaluatin ties.

Finally, the Spearman rank correlation coefficient was used to test the null hypothesis that there was no association between the evaluation methods. The null hypothesis was rejected in every case. Based on these results, it is the recommendation of this study that formal application of multi attribute utility assessment techniqeus be applied to generate a substitute for the signature table evaluation technique applied in this case.

DATE
ILMED
-8